

UC San Diego

UC San Diego Previously Published Works

Title

Kant on materialism

Permalink

<https://escholarship.org/uc/item/90p8s4b3>

Journal

British Journal for the History of Philosophy, 24(5)

ISSN

0960-8788

Author

Watkins, Eric

Publication Date

2016-09-02

DOI

10.1080/09608788.2016.1230091

Peer reviewed

Kant on Materialism

Abstract: In this paper I argue that Kant's complex argument against materialism involves not only his generic commitment to the existence of non-spatio-temporal and thus non-material things in themselves (which follows directly from Transcendental Idealism), but also considerations pertaining to reason and the subject of our thoughts. Specifically, I argue that because Kant conceives of reason in such a way that it demands a commitment to the existence of the unconditioned so that we can account for whatever conditioned objects we encounter in experience, our thoughts, which are also conditioned, require something unconditioned that, because it is unconditioned, cannot be material. In this way, Kant's attitude towards materialism is based not only on abstract features of his metaphysics and epistemology, but also on specific features that were under serious discussion in the early modern period.

Keywords: Kant, materialism, unconditioned, I, psychology

Kant's views on materialism in the *Critique of Pure Reason* have long been a source of puzzlement. On the one hand, refuting materialism seems to be one of the goals that motivates his overall project. To preserve the possibility of moral obligation, moral action, and moral responsibility, freedom must be defended against all objections, and one of the most dangerous threats to freedom was thought to be materialism, given that a materialist conception of human beings seemed to leave no room for the possibility of freedom insofar as the behavior of matter is fully determined by mechanistic laws. In line with this thought, Kant claims in the second edition Preface to the first *Critique* that he intends to 'sever the very root of materialism' (Bxxxiv).¹ On the other hand, it is striking that in a work that extends over eight hundred pages, very little space is explicitly devoted to the topic, and when it is discussed, it is mentioned primarily as an almost accidental corollary to the Second Paralogism's treatment of the soul's simplicity (B420).² Given the threat that materialism seemed to represent to freedom at the time and

given Kant's recognition of the severity of the threat, it is surprising that he did not devote more explicit discussion to the topic.

Now Karl Ameriks has helped to remove some of the mystery here by showing that Kant takes materialism, as a general thesis, to be ruled out by Transcendental Idealism. Transcendental Idealism is famously based on a distinction between appearances and things in themselves, where appearances are spatio-temporal objects that depend on our subjective forms of intuition, and things in themselves are non-spatio-temporal objects that exist independently of our subjective forms of intuition.³ Ameriks brings Transcendental Idealism to bear on materialism by way of the following argument: If matter is defined as the movable in space, then it is essentially spatial but if, as Transcendental Idealism maintains, things in themselves are not spatial, then things in themselves cannot be matter.⁴ Given Kant's commitment to the existence of things in themselves, there must be things that are not matter. Now this is surely a good refutation of materialism, as far as it goes, and it is also plausible to hold that this is at least part of Kant's justification for claiming that he is in a position to refute materialism.⁵

At the same time, even if Ameriks' account is correct, it may not be the whole story. For if it turned out that things in themselves were indeed immaterial and that all appearances consisted entirely of matter, the materialist might well concede to having lost a minor battle (regarding things in themselves), but declare a resounding victory in the larger war.⁶ For the materialist could, in a first step, distinguish between different kinds of

materialism, with universal materialism being the doctrine that *everything* consists only of matter, cosmological materialism that everything *in the world* consists only of matter, and psychological materialism that *human beings* consist only of matter.⁷ And the materialist could then, in a second step, note that even if one were to concede that Kant had refuted the first two kinds of materialism, that would still not immediately entail the falsity of the last kind, for even if some things in the world are not material, human beings might still be material through and through.⁸ That is, even if one were to accept Transcendental Idealism and that every (spatio-temporal) appearance must be grounded in a (non-spatio-temporal and thus non-material) thing in itself, the psychological materialist could still maintain that a human being is exclusively material, even while granting that human beings are grounded in something that is not material. What's more, it is clear that psychological materialism is especially relevant in this context since what most interested both early modern materialists and their opponents, including Kant, was the feasibility of providing explanations of human beings that appeal only to matter and the mechanistic laws of motion that govern it.⁹ Being told that unknowable things in themselves, which (somehow) ground the spatio-temporal world of appearances, are not material, might be viewed as an instance of 'spooky metaphysics' that materialists would hardly embrace with open arms, but one could imagine that they might simply dismiss that point as basically irrelevant to their

primary concern with explaining what human beings are and how they function in the material world.

To find an interpretation of Kant's refutation of materialism that advances an argument that can supplement the one Ameriks has already presented, we must focus on those features of his position that are specific to human thought and on how those features can be exploited against the psychological materialist. It turns out that Kant's most promising argument against the psychological materialist draws on features that emerge from his distinctive analysis of reason as a spontaneous faculty that searches for conditions. More specifically, we must understand (1) Kant's account of real (as opposed to logical) conditions, (2) how, in the first three Paralogisms, representations require different kinds of real conditions, and (3) how reason can make use of real conditions in inferring from the existence of conditioned objects to the existence of something unconditioned.¹⁰ With a clearer understanding of these three points and the larger picture that they paint, we can turn to Kant's analysis of the Paralogisms and, on that basis, reconstruct his argument against psychological materialism. The basic idea will be that the I cannot, according to Kant, be material because what is unconditioned cannot be material and the I, whatever its ontological status it might be, is the unconditioned condition of my thoughts. By pursuing this line of thought, we can see how it is that Kant hopes to refute the psychological materialist and thereby remove one significant threat to freedom.

I. Reason, Real Conditions, and the Unconditioned

It is a distinctive feature of Kant's account of reason that it is defined as a spontaneous faculty that seeks not only the *conditions* for what is conditioned, but also the *totality* of conditions and hence the *unconditioned*. This point can be expressed in more contemporary terms, and perhaps less obscurely, by noting that reason is interested not only in explanations (conditions) of whatever stands in need of explanation (what is conditioned), but also in total and complete explanations, that is, in explanations that cover everything that stands in need of explanation (i.e., that explain everything conditioned) and in terms that do not themselves admit of further explanation (i.e., in what is "unconditioned"). For only then will reason have found a satisfactory 'resting place' (A584/B612). We will return to the unconditioned shortly, but we must first understand his conception of a condition more fully, especially since the unconditioned has to be explained in terms of it.

Though Kant is fully committed to the unity of reason—reason is one and the same faculty in all contexts in which it is employed—he distinguishes between the logical and real uses of reason (A303/B359, A305/B362). Reason in its logical use gives rise to syllogisms, which express, Kant maintains, logical conditioning relations between cognitions (namely between the major and minor premises and the conclusion). Specifically, the major premise expresses a condition, and the minor premise is subsumed

under the condition expressed in the major in such a way that the conclusion then logically follows. Given this conception of a syllogism, it makes sense to say that the premises are conditions for the conclusion, which is conditioned. What's more, Kant thinks that the different kinds of syllogisms—categorical, hypothetical (e.g., modus ponens and tollens), and disjunctive (A304/B361)—are determined by the different kinds of conditioning relations in the major premise (in the form of categorical, hypothetical, and disjunctive judgments). Thus, by way of the notion of a logical conditioning relation, which holds between cognitions, Kant is able to bring his characterization of reason together with his theory of syllogisms.

But Kant also identifies a real use of reason, which pertains not to *logical* conditioning relations between *cognitions* in syllogisms, but rather to *real* conditioning relations between *objects* (taking 'object' here in a very broad sense so as to include both physical and mental entities). That is, just as a conclusion depends logically on the major and minor premises in a syllogism, so too an object (or its state) can depend on other objects in the world. The basic idea here is that the notion of a real condition gives expression to what we might call metaphysical dependence. And just as Kant is committed to different kinds of logical conditions in different kinds of syllogisms, so too he is committed to different kinds of real conditions between different kinds of objects. For example, Kant clearly holds that (1) one moment in time is a condition of the moment that follows it, (2) parts are conditions of the whole they compose, and (3) a cause is a condition of its

effect. Given these and other examples, Kant seems to be operating with a generic notion of real conditioning that involves an asymmetrical, transitive, and intelligible relation of metaphysical dependence, and these three examples are all different specific instances of that generic notion.¹¹

In the Transcendental Dialectic, Kant makes various remarks that clarify the different kinds of real conditions that he acknowledges. The three examples just listed are discussed, for instance, in the (first three) Antinomies, which are interested in the ‘unity of objective conditions in appearance’ (A406/B432), that is, in the different kinds of real conditions of objects that appear in the (spatio-temporal) world. The Ideal of Pure Reason, by contrast, is interested in the ‘unity of objective conditions of the possibility of objects in general’ (A406/B433), that is, in the real condition of all possibility, while the Paralogisms concern ‘the subjective conditions of all representations in general (of the subject or the soul)’ (A406/B433), that is, in the real conditions of all of our representations.

Now one can see, in a general way, how the Paralogisms focus on the real conditions of our representations, since representations are obviously conditioned in various respects. For representations clearly depend in different ways on a mental subject, which can be expressed by ‘I’ or ‘the soul’ insofar as they must be *had*, or *thought* by a mental subject and are thus dependent (conditioned) on such a subject.¹² To put the point in terms closer to Kant’s own, one can say that the I referred to in the phrase ‘I think’, which is ‘the sole text’ of the Paralogisms (A343/B401), is a real condition of

all representation. For representations are conditioned by the “I think”, which is in turn conditioned by the I that makes “I think” possible insofar as it thinks (i.e. engages in the activity of thinking).

Further, since each of the four Antinomies has a different real conditioning relation as its subject matter and since there are four distinct Paralogisms, it stands to reason that each Paralogism focuses on a different real conditioning relation.¹³ That is, each Paralogism concerns a different feature of representations that is conditioned and considers what feature of the mental subject is required to explain that conditioned feature. For our purposes, it will suffice to understand the specific real conditioning relations for the first three Paralogisms.

The First Paralogism notes that representations are not free-floating, metaphysically independent entities that could somehow think themselves and thus stand on their own. Instead, as representations, they must be mental states *of* a subject, in the sense that they must *inhere in* the mental subject that thinks them. But for representations to inhere in the mental subject that thinks them, that mental subject must, it seems, be a substance, given that accidents must ultimately inhere in a substance and only a substance can act in the way required for the acts of thinking that allow representations to inhere in a substance.¹⁴ Thus, the I insofar as it is a substance that thinks representations is a real condition of the representations that inhere in it.

The Second Paralogism focuses on the fact that a representation is not a collection, or aggregate, of parts, but rather possesses a special kind of unity that (e.g. bodily) composites do not have. This (mental) unity can be explained, however, only if the mental subject that thinks that representation is simple. For the unity of a representation requires a single act of thinking (rather than a collection of distinct acts), which is possible, it seems, only if the mental subject that performs that act is simple. Thus, the I insofar as it is simple in its act of thinking is a real condition of the representations that the I thereby thinks.

The Third Paralogism picks up on the fact that a plurality of representations can be attributed to a single, identical mental subject. This point has several different aspects. First, it is not necessary that one posit a distinct mental subject for each numerically distinct representation, because different representations can be attributed to the same mental subject. Though Kant formulates his point with respect to personal identity over time—since he is interested in the human case, which is temporal—it can also be put more generally, since God can be an identical mental subject of a plurality of representations without involving personal identity over time (given that God, for Kant, cannot be temporal). Further, representations are not ownerless; instead, they must belong to, and be attributed to, a person, who, in thinking the representation, is thereby responsible for that representation belonging to it.¹⁵ Finally, in light of these two points, it follows that a plurality of representations can be attributed to a single identical

person insofar as that person actually thinks all of these thoughts. Thus, the I insofar as it is a numerically identical person that thinks is a real condition of the plurality of distinct representations that can all be ascribed to it.

Accordingly, the first three Paralogisms are concerned with how different features of the I—its substantiality, simplicity, and personal identity—are different real conditions of different conditioned features of representations—their accidentality, unity, and attributability—if the I is thinking these representations.

As noted above, Kant understands reason as a faculty that seeks not only the conditions of whatever is conditioned, but also the totality of such conditions and thus the unconditioned. This connection of reason and the unconditioned plays out in two main ways for Kant, one concerning our representations of the objects of traditional metaphysics, the other concerning a line of argument that would establish the existence of something unconditioned. Consider first the connection that Kant draws between our ideas of reason and the objects of traditional metaphysics.¹⁶ An idea, for Kant, is a representation ‘to which no congruent object can be given in the senses’ (A327/B383). Since reason is the faculty that seeks the unconditioned, an idea of reason is a representation of an unconditioned object that cannot be given in the senses. Given these views, Kant then goes on to claim that the objects of traditional metaphysics—God, freedom, and the soul—are represented by means of ideas of reason. For these objects cannot, Kant maintains, be given in the senses and they are all

unconditioned in some respect. For, according to the Judeo-Christian tradition, God would be properly characterized as unconditioned insofar as nothing is the cause, or condition, of his existence. As Kant understands (transcendental) freedom, a free action would also be properly characterized as unconditioned, because a free action cannot be caused, or conditioned, and still be free. Finally, because the soul is the ultimate subject of our representations—nothing other than, or lying behind, the I could think our representations—it, too, is unconditioned in this respect.¹⁷ Thus, reason is the faculty that represents the objects of traditional metaphysics by means of its ideas.

Second, in the context of his discussion of traditional metaphysics, Kant highlights a line of argument with respect to the unconditioned that he ascribes to reason.¹⁸ The basic idea is that if the conditioned is given, then the totality of conditions must also be given, since without all of its conditions, the conditioned would not be given. But if the totality of conditions is given, then the unconditioned must be given too, for the totality of conditions entails the unconditioned.¹⁹ Kant repeats this argument, or at least its basic premises, on numerous occasions. For example, in the Introduction to the Transcendental Dialectic, he remarks: ‘if the conditioned is given, then so too is the whole series of conditions subordinated to each other, which is itself therefore unconditioned (i.e., contained in the object and its connection)’ (A307-8/B364). In the Antinomy of Pure Reason, he similarly remarks: ‘Reason demands this in accordance with the principle: **If**

the conditioned is given, then the whole sum of conditions, and hence the absolutely unconditioned, is also given, through which alone the conditioned was possible' (A409/B436). Shortly thereafter, he reiterates: 'The absolute whole of the series of conditions for a given conditioned is always unconditioned, because outside it there are no more conditions regarding which it could be conditioned' (A417-8/B445).

Kant's most basic line of thought can, I believe, be reconstructed roughly as follows. Take any given object *x*, which is conditioned.²⁰ Since *x* is conditioned, it is an analytic truth that there must be at least one or more *ys* that happen to condition it. Otherwise, it would not be conditioned. Now let *t* be the sum of all *ys* that condition *x* ('the whole sum of conditions' or the 'absolute whole of the series of conditions'). It immediately follows that *t* must be unconditioned. For suppose, for the sake of a reductio, that *t* is conditioned. If *t* is conditioned, then it is an analytic truth that there must be some *z* that conditions it. Given transitivity, if *z* conditions *t* (by the reductio's original assumption) and *t* conditions *x* (by being the sum of all *ys* that condition *x*), then *z* conditions *x*. But then, contrary to the definition of *t*, *t* does not in fact contain all of the conditions of *x*. Therefore, *t* cannot be conditioned, i.e. something unconditioned must exist. In short, as soon as the existence of something conditioned is granted, reason is committed to the existence of the unconditioned as well.

Now it is crucial to note the exact status of the conclusion of this argument, lest one mistakenly succumb to the kind of 'transcendental

illusion' that Kant explicitly warns against in the Transcendental Dialectic. That is, the argument claims that, given the existence of a conditioned object, reason is committed to, or must posit, the existence of something unconditioned, which conditions the object and thus can explain its status as conditioned. It does *not* follow that we must be able to have *cognition* of the existence of the unconditioned. For to have cognition of an object, that object would have to be both given in sensibility and thought through the understanding, and Kant thinks that unconditioned objects cannot be given to us through the senses (A308/B365, A311/B367, A327/B383, A483/B511, A531/B559). Though Kant's characterizations of transcendental illusion are vague and subject to different interpretations, one kind of 'natural and unavoidable' illusion (A298/B354) that he could be pointing to is that of accepting the following closure principle (which is false): because one cognizes x and reason can validly infer from x to y, one can, for that reason, cognize y. That is, just because reason's inferences are truth-preserving, it does not follow that they are therefore also cognition-preserving. So, although reason is *right* to infer from the existence of the conditioned to the existence of the unconditioned, it would be a mistake to infer that we must therefore be able to have *cognition* of the unconditioned.

What's more, this point is absolutely central to Kant's critique of traditional metaphysics. For traditional metaphysics, on Kant's analysis, claims to have cognition of unconditioned objects that is both synthetic and a priori, that is, substantive and yet not based on particular experiences.

What the argument provided above shows is that it is possible to *infer* to the existence of something unconditioned without relying on experience of that object. However, it is not possible to extend our *cognition* so that it would deliver substantive new results about unconditioned objects, such as that we are free or that God exists. Instead, all that we are rationally committed to is that something unconditioned exists (and that it is a condition of the conditioned objects that we do cognize), but we cannot cognize (or even infer) any of its positive intrinsic features.²¹ In fact, even when we consider an object not merely as conditioned, but as conditioned in some specific respect and pursue the real conditioning relations back to the particular unconditioned object that conditions it, we still are unable to cognize any of its positive intrinsic features, much less that it must be, e.g., God (as a being endowed with understanding and will). Kant testifies to the indeterminacy of this commitment, when he explicitly remarks: ‘And now we are thinking of a Something about which we have no concept at all of how it is in itself, but about which we think a relation to the sum total of appearances [...]. If, accordingly, we assume such ideal entities, then we do not really extend our cognition beyond the objects of possible experience’ (B702). That is, we view the existence of the unconditioned that we assume, or posit, so as to have a real unconditioned condition of the conditioned objects that we cognize, as merely ‘a Something’, since we cannot attribute new positive determinate predicates to it, thereby thwarting the ultimate goal of traditional metaphysics.²²

Even though we cannot extend our cognition by adding positive new intrinsic determinations to the unconditioned, the kind of ontological commitment that we have toward the unconditioned (insofar as we are rational) fits nicely with Kant's claim that ideas of reason function as regulative principles. For this argument shows that reason is committed to the existence of something unconditioned and reason's ideas, which represent unconditioned objects, function as regulative principles that direct us to pursue these objects as far as we can, even if we cannot cognize them. Though our cognitive limitations prevent us from satisfying reason's desire to obtain cognition of an unconditioned condition, Kant's argument shows that reason is at least committed to the existence of what our regulative principles of reason demand that we necessarily seek.

What has emerged from consideration of Kant's account of reason, the notions of a condition it employs, and its commitment to the existence of the unconditioned, is a larger picture of how reason functions. Reason seeks not only logical conditions in syllogisms, but also real conditions of objects that are given to us, where there are different kinds of real conditions for different kinds of objects. Here we have focused on the different real conditions of our representations, which revealed different features that the I must have to be the real condition of the different conditioned features of representations of which we are aware. Further, reason is committed to the existence of something unconditioned, though reason cannot generate the kind of synthetic a priori cognition that was the goal of traditional metaphysicians

prior to Kant, despite the natural illusion to think that we can. Instead, metaphysics proceeds by forming ideas of unconditioned objects that regulate our intellectual inquiry by helping us to find as many real conditions as we can, and even though we cannot achieve cognition of the unconditioned, we can get ever closer and thus approximate such cognition. In this way, metaphysics does not achieve the kind of cognition it had hoped for, but it does provide a crucial service to our intellectual engagement with the world by revealing as much rational structure in the world as we are capable of discerning.

II. The Paralogisms and Materialism

With this broader picture of reason and the role it plays in traditional metaphysics in mind, we can now return to Kant's most explicit discussion of materialism in the first *Critique* in the Paralogisms. Kant's view is complicated by the fact that he explicitly rejects those arguments in the (First and Second) Paralogisms that would contribute most directly to establishing the soul's immateriality (and thus the falsity of materialism). In light of his rejection of these arguments, any argument he might endorse that would refute materialism will have to differ from these arguments in some way. As a result, we first have to understand the basic structure of the first two Paralogisms as well as Kant's complex diagnosis of the mistakes that their arguments make. Only then can we see what resources are still available to construct an argument against materialism that differs from the

arguments he rejects as paralogistic. In this way, we will be able to see how he could be justified after all in claiming that he has ‘severed the very root of materialism’ and thereby diffused an important threat to the possibility of freedom.

In the Paralogisms, Kant considers four arguments—categorical syllogisms that are ascribed to ‘rational psychologists’, who take nothing more than ‘I think’ as their premise—that have as their conclusion that the soul has a certain property. The argument under consideration in the First Paralogism concludes that the soul is a substance and that in the Second infers to the soul’s simplicity. In both cases, the major premise states the meaning of a particular term that expresses a defining feature of that property (substantiality and simplicity). The minor premise then notes that the I has the feature that figures in the definition of each property. The conclusion then attributes the property defined in the major premise to the soul. On the basis of these arguments, the rational psychologist then claims to have cognition that the soul is a simple substance. Since the arguments for this substantive conclusion depend only on (1) a definition and (2) aspects that are revealed through ‘I think’ (rather than any particular experience), the cognition that is the conclusion of the syllogism would be both synthetic and a priori, which, as we have seen, fits Kant’s characterization of the status of the claims of traditional metaphysics. And although immateriality is not the official topic of the First or Second Paralogism, it is closely related to them insofar as claiming that the soul is a

simple substance forms the crucial premise in an argument showing that it must also be immaterial.²³ Thus, if the arguments attributed to the rational psychologist in the First and Second Paralogisms were sound, materialism would be in serious danger.

However, as noted above, Kant rejects these arguments, offering a subtle and complex diagnosis of the errors that they commit.²⁴ First, he shows how the argument can be interpreted to be either invalid due to an ambiguity in the middle term or else valid, but then one whose conclusion must be understood as trivially true and thus of no use for the metaphysician who wants to draw substantive implications from it (such as that the soul must also be immortal). Second, Kant shows that it would be a mistake to take the major and minor premises to be analytic and the conclusion synthetic. Third, it would be problematic to move from premises that concern how we must *think* of the soul to a conclusion about how the soul must *be*, an error he points out more clearly in the B-edition. Fourth, Kant takes care to distinguish what can legitimately be said about the self on the basis of his own positive doctrines, such as the unity of apperception (which involves 'I' in a constitutive way), from the conclusions that the rational psychologists want to draw merely on the basis of an analysis of the concept 'I think'.²⁵ While there are still further aspects to Kant's diagnosis of how the different arguments go wrong, I take it that the most fundamental problem, on which many of the others depend, is that the soul is not given to us in intuition and since cognition requires that the object of cognition be given in intuition, the

soul cannot be cognized.²⁶ Thus, what many of the diagnoses are pointing out in different ways is simply that no argument that would purport to deliver cognition of the soul based purely on a priori considerations can justify any substantive conclusion in the absence of an intuition of the soul.

Given this general diagnosis, several salient points about the First and Second Paralogisms stand out. As for the first Paralogism, one strand of Kant's diagnosis reveals that one cannot have cognition of the I as a substance because, as Hume astutely noted (albeit in different terms), we do not have an intuition of a permanent object to which the category of substance could be applied so as to generate cognition. When I look inside myself, I do not see a substance, certainly not in the same way that I see tables and chairs (which, as configurations of matter, are spatial substances). Now it is true (perhaps now contra Hume) that one can be aware of some kind of subject, an I, that can always accompany my representations (e.g. by being the I that is presupposed by all of our judgments), but Kant notes that the representation of such an I is devoid of content and thus does not warrant any *determinate* or *specific* ontological claims about the nature of the I that is the subject of our representations (e.g., that it is a permanent substance). Instead, the I that is presupposed as the subject of our judgments is what Kant refers to as a 'logical subject' and is a substance only in a different and quite limited sense, namely that of a 'substance in concept' (A400), which 'is to be represented as a subject in itself without in turn being the predicate of another subject' (A401).²⁷

However, and crucially, such a logical subject does not suffice for cognition that the I must be a substance in any more robust sense, for, now drawing on a different strand of Kant's diagnosis, even if one acknowledges that we must *think* that there is a logical subject that underlies all possible judgments, it does not follow that we can determine, or cognize, what kind of ontological entity such a logical subject must actually *be*. That is, what *instantiates* that logical subject could, for all we know, be an accident of a more fundamental substance. As a result, Kant's general diagnosis of the errors of the Paralogisms helps us to understand why the argument of the First Paralogism cannot justify cognition that the I that is in fact the subject of all of my representations must be a substance.

Analogous considerations are applicable to the argument at issue in the Second Paralogism. Since we have no intuition of the I as a distinct object within ourselves, we also have no intuition of its simplicity and thus no cognition that the I is simple. And even though the I that is presupposed by all of our judgments (or representations in general) is, in some sense, simple (i.e. it lacks parts), one might think of it as a purely formal feature of the subject of all judgments, since it is devoid of the content that would be needed to allow for cognition of its simplicity. What's more, even if we must *think* that all of our judgments presuppose such a logical subject that is, in some sense, simple, we still cannot cognize that what *realizes* the logical subject must be ontologically simple, rather than complex. This last objection is particularly relevant to the materialist, since a materialist might well

concede that thought requires a substance, but then deny that it requires a *simple* substance, since the brain is a complex entity with various parts whose proper functioning is required for thought to occur. Thus, in the Second Paralogism just as in the First, serious objections to the argument prevent us from having synthetic cognition of the ontological nature of the I that thinks.

Now one might think that with the failure of the arguments at issue in the First and Second Paralogisms, there is no longer any hope of finding a justification for Kant's rejection of materialism. With what resources could one accomplish such a task, if his diagnoses of the relevant arguments in the Paralogisms have been successful in undercutting the most tempting inferences to substantive features of the soul? However, it is important to keep two points in mind here. First, note that Kant's arguments target the possibility of *cognition* of substantive features of the self, where cognition is a mental state that must satisfy two non-trivial conditions, namely that the object of cognition must be given in sensible intuition and thought through concepts. And in fact, the failure to satisfy the first of these two conditions was the fundamental problem that Kant identified with the relevant arguments in the Paralogisms. Yet the fact that we cannot have cognition of substantive features of the soul does not immediately entail that we cannot infer that these substantive features obtain (even if we cannot cognize how they would do so). Granted, any inference that one wants to draw must be

justified, and that is no simple task, but the logical space in which one could *try* to develop such an argument is nonetheless available.

Second, as we have seen above, on Kant's account, reason is interested not only in logical conditions that obtain between the premises and the conclusion in a syllogism, but also in the real conditions that obtain between different objects in the world, for reason seeks to infer from the existence of conditioned objects to the existence of their unconditioned condition. As we also saw above, this is true for our representations as well as for other kinds of objects, since representations are conditioned in different respects and require different kinds of real conditions, for without these real conditions obtaining, the representations could not exist at all. What's more, Kant clearly indicates that he views the subject matter of the Paralogisms in precisely these terms, for he notes: 'Pure reason is concerned merely with the absolute totality of this synthesis, i.e., with that condition that is itself unconditioned [...]. Because, further, the only condition accompanying all thinking is the I, in the universal proposition 'I think', reason has to do with this condition insofar as it is itself unconditioned' (A397-8). Thus Kant clearly states that the I that reason is interested in is the unconditioned condition of all thinking or all representations.

Taking these two points together, we can now formulate what I take to be Kant's most compelling argument against the psychological materialist. The argument assumes that a representation is a conditioned object that requires a real condition; thoughts do not think themselves, but rather

require an I that serves as their real condition in actively thinking them. As we saw above, this basic fact plays out along several dimensions.

Representations are not self-subsistent, but rather accidents; they are not an aggregate of parts, but rather have a mental unity; they are not owner-less, but are rather attributable to a logical subject, or person. But at this point, Kant's argument to the unconditioned becomes relevant, for, as we have seen, reason is justified in inferring from the existence of something conditioned not only to its conditions, but also to the existence of something unconditioned (namely the unconditioned condition of the conditioned representation). As a result, the existence of a representation entails the existence of something unconditioned.

But what is the unconditioned something that is the real condition of representations? As we saw above, Kant's generic account of reason (in terms of conditions in general) is unable to offer an informative answer to this question. However, our analysis of the *real* conditions of *representations* reveals that the I that is the logical subject of my representations cannot be conditioned by anything else, because the I that thinks my thoughts cannot be backed up, or supported, by a further I in its thinking; no one else can think my thoughts for me, and nothing else can bestow unity on my representations.²⁸ Because the logical subject that is the real condition of my representations cannot be conditioned further, it must be, as we have seen, unconditioned.²⁹

However, we can now invoke a more familiar line of thought. Since Kant maintains that the unconditioned cannot be given in sensible intuition, which is in turn required for something to be an appearance, it follows that whatever is unconditioned cannot be an appearance. But if, as we saw above, the logical subject is unconditioned, then it too cannot be an appearance.³⁰ And, if, according to Transcendental Idealism, matter is essentially an appearance (due to its spatiality), then the logical subject, or I that thinks, cannot be matter either. Therefore, psychological materialism, which claims that human beings consist entirely of matter, must be false.

Note that this argument is fully compatible with Kant's diagnosis of the problems that plagued the arguments at issue in the Paralogisms.³¹ For those arguments inferred, mistakenly, that we can have cognition of the soul as a simple immaterial substance. The argument presented above does not claim that we can have cognition that the soul is immaterial. Instead, reason infers that the I that is the logical subject cannot be material, and the inference is based on reason alone, and not on any intuition of the soul. As a result, it does not allege that we have cognition of the soul. Further, the argument does not infer from how we must think of the logical subject to how the soul must be. For example, it does not infer that the logical subject is a substance (in its typical robust sense), since it does not speculate as to what the ontological realization of the logical subject is.³² The logical subject could be a substance in the appropriate sense, but it could also, for all we know, be an accident of some other substance.³³ Finally, the argument does not ascribe

any positive intrinsic features to the logical subject that would go beyond what is required by the real conditioning relations involved in (conditioned) representations that form the basis of the argument. Reason does infer that the logical subject cannot be material, but that is a negative claim, one that does not entail any positive characterization of what the logical subject (or its ontological realization) must be.

The argument is also able to fill precisely the gap that Ameriks' position left open. The argument Ameriks attributes to Kant is successful at showing that universal and cosmological materialism are untenable, but does not have the resources to rule out psychological materialism, because the considerations it is based on are entirely generic, pertaining to all things in themselves. The considerations that drive the argument presented above, by contrast, turn on specific features of representations, namely the particular ways in which they are conditioned and the particular ways in which the logical subject can serve as a real condition of those representations. As a result, the argument presented above is able to show that the subject of our thoughts, which is an intimate part of who we are as human beings, cannot be material, and thus that we cannot be purely material beings, as the psychological materialist claims.

III. Conclusion

In this way, Kant can claim to have refuted the materialist as an ontological position on the nature of the soul, or of the I that thinks.

However, Kant's position turns out to be quite nuanced in a way that allows for the possibility that he is not as far from the materialist position in certain respects as one would initially think. Though Kant is committed to whatever is unconditioned being immaterial, he does not think that we can cognize whatever it is that is the subject of our thoughts. But we must still seek explanations for whatever empirical representations we have, given that reason's idea of the immaterial soul serves as a regulative principle for explaining our representations; but since reason is interested in obtaining as many cognitions of psychological conditions as possible and yet cannot obtain cognition of anything unconditioned, it turns out that for any cognition that we might attain, both the *explandandum* and the *explanans* will have to be conditioned. Thus, when it comes to providing particular explanations of particular thoughts, Kant will never be able to take refuge in his immaterialist ontological commitments. In that stance he is every bit as committed as the materialist to providing detailed *empirical* explanations (even if not exclusively materialist ones).

At the same time, this is not tantamount to saying that Kant is a materialist when it comes to the nature of what thinks or to providing psychological explanations thereof. Indeed, on the latter point, Kant was committed to what Ameriks has called scientific immaterialism throughout much of his career. For the empirical conditions of our representations are not explicable by way of the laws that govern matter. In this respect, though Kant is in principle open-minded about the possibility of scientific

materialism, he does think that, as the science of matter happens to work out, immaterialism seems to be the case.³⁴ But note that this position does not follow immediately from features of his broader philosophical system, but rather from his assessment of the adequacy of psychological scientific explanation in terms of mechanistic laws.

As a result, we can now better appreciate the place of the debate about materialism within Kant's system. While Kant's Transcendental Idealism does entail the falsity of universal and cosmological materialism, since there must be at least some immaterial things (things in themselves), he also deploys more specific resources in the course of articulating why the ultimate subject of our thoughts must also be immaterial. Because the I that is the real condition of the representations that are ascribed to it must be unconditioned, it cannot be a material thing. Thus, there is in the end a proper justification for his claim that he intends to sever the root of materialism, where one can take materialism not merely as a global metaphysical thesis, but also as one that applies to the specific subject matter of human beings, which was at the heart of debates throughout the seventeenth and eighteenth centuries.³⁵

BIBLIOGRAPHY

Allison, Henry. 1996. 'Kant's Refutation of Materialism,' in *Idealism and Freedom*, Cambridge: Cambridge University Press.

Ameriks, Karl. 2001. *Kant's Theory of Mind*. Oxford: Clarendon Press.

Chignell, Andrew. 2007. 'Belief in Kant,' in *The Philosophical Review* 116 No. 3: 323-60.

Dyck, Corey. 2013. *Kant and Rational Psychology*. Oxford: Oxford University Press.

Kant, Immanuel. 1998. *Critique of Pure Reason*. Translated and edited by Paul

Guyer and Allen Wood. Cambridge: Cambridge University Press.

Grier, Michelle. 2001. *Kant's Doctrine of Transcendental Illusion*. Cambridge: Cambridge University Press.

Proops, Ian. 2010. 'Kant's First Paralogism'. *The Philosophical Review* 119, No. 4: 449-495.

Rosefeldt, Tobias. 2000. *Das Logische Ich. Kant über den Gehalt des Begriffes von sich selbst*. Berlin: PHILO.

Rumore, Paola. 2014. 'Kant's understanding of the Enlightenment with reference to his refutation of materialism,' in *Con-Textos kantianos* No. 1: 81-97.

Watkins, Eric. forthcoming. 'Kant on Real Conditions,' in (ed.) Violetta Waibel and Margit Ruffing. *Akten des 12. Internationalen Kant-Kongresses, 'Natur und Freiheit' in Wien vom 21.-25. September 2015*, Berlin.

Wunderlich, Falk. 2001. 'Kant's Second Paralogism in Context' in (ed.)
Wolfgang Lefevre. *Between Leibniz, Newton, and Kant: Philosophy and
Science in the Eighteenth Century*. Dordrecht: Kluwer.

¹ Citations to the *Critique of Pure Reason* are to the standard (first) A and (second) B editions of 1781 and 1787. All translations are my own, though I have consulted the Cambridge edition translations.

² In this paper, I abstract both from Kant's practical considerations as well as from works other than the *Critique of Pure Reason*.

³ See esp. the Preface and the Postscript to the second, revised edition of Ameriks (2000).

⁴ For a statement of Ameriks' argument, see (2000, 309). Ameriks notes that his position is general (i.e., not a specific consequence of rational psychology) as follows: 'Kant's immaterialism does not rest essentially on claims specifically about the mind' (2000, 308).

⁵ In transcripts from his metaphysics lectures Kant repeatedly distinguishes matter and material. Monads, which are simple, cannot be matter, but they are material if they constitute matter. I abstract from Kant's terminology in the following.

⁶ To call something immaterial is not to assert anything positive of that thing, but rather simply to say that it is not material. Specifically, it is not to assert that things are in some sense 'minded', or mental. Kant is explicit about the difference between immaterialism and spiritualism (which is committed to the mind necessarily being able to exist without a body), noting that both materialism and spiritualism are unacceptable extremes (B420). Cf. Ameriks (2000, xvii).

⁷ Now Ameriks might respond by noting that there must be a non-material thing in itself underlying human beings, and not simply other things in the world, and that therefore even human beings cannot be material. However, unless one can show that there is some significant aspect of human beings that

is not material, this response has no bite. So what is needed is to show that some substantive aspect of being a human being cannot be entirely material. (There can be reasonable disagreement about what exactly counts as 'substantive' as well as about what Kant's ultimate interests in human beings are.) – Note also that one could enrich these different kinds of materialism by adding that everything, the world, or human beings must function exclusively according to mechanistic principles.

⁸ One way to deny that everything consists entirely of matter is to assert that one thing consists of something other than matter, which is still consistent with everything else consisting of matter. Ameriks takes Kant's doctrine of immaterialism to be stronger than simply the falsity of universal materialism, for he thinks that 'every concrete thing is basically non-material' (2000, 308).

⁹ For discussion of the distinction between the different kinds of materialism in Germany at the time, see Rumore (2014).

¹⁰ This fits with Kant's discussion of materialism in the Second Paralogism.

¹¹ See Watkins (2018).

¹² I take it that Kant is interested here in representations that are something for me (as opposed to those that might simply be in me), that is, in representations to which I can add 'I think'.

¹³ This interpretation is confirmed by the table that Kant lays out at A344/B402.

¹⁴ It is true that an accident could inhere in a further accident, but, rightly or wrongly, Kant does not accept that there could be accidents "all the way down". Thus he thinks that the existence of an accident ultimately requires the existence of a substance. We return to this issue below.

¹⁵ Kant conceives of a person (as opposed to a thing) as something to whom actions are attributable. Insofar as a person is merely a mental subject, that

person's representations must be attributed to her in the sense that the person is mentally responsible for those representations. Insofar as a person is an agent, the person's actions must be attributable to her in the sense that the person is, depending on the nature of the action, prudentially and morally responsible for those actions.

¹⁶ Kant develops his account of the regulative principles of reason in detail in the Appendix to the Transcendental Dialectic.

¹⁷ See Kant's attempt in the second and third sections of the First Book of the Transcendental Dialectic, 'On the Concepts of Pure Reason'.

¹⁸ I discuss this argument in detail in a longer, currently unpublished manuscript.

¹⁹ It is clear, e.g., from Kant's discussion in the Antinomies at A498/B526, that "given" can have (at least) two different meanings, depending on whether things in themselves or appearances are at issue. For things in themselves to be given means that they exist, whereas for appearances to be given can also mean that the subject is aware of them (or their existence) as represented in sensible intuition. But note that Kant accepts the conclusion of the argument regardless of whether what is given are things in themselves or appearances. Now one might agree that the existence of a conditioned thing in itself does in fact entail the existence of an unconditioned thing in itself, but then argue that for appearances, the existence of a conditioned appearance entails only that a regress in the series of all of its conditions and thus the unconditioned be given as a task (*aufgegeben*), not that anything unconditioned exist. However, it is crucial to understand exactly what the regress that is given as a task involves and how logical and real conditions apply to appearances. If the task that is given to us concerns the logical conditions of the cognition of the appearance

that is given in sensible intuition, then the logical use of reason demands that we find the premises from which that cognition follows, but since we may not already be in possession of cognitions that can serve as premises for the requisite syllogism, finding cognitions that can serve as those premises is given to us as a task. If, however, we are interested not in the logical conditions of the cognition of the appearance that is given to us, but rather in the real conditions of that appearance, then it is true that we must seek cognition of its real conditions, but also that its real conditions must exist (though not necessarily as an appearance), for if its real conditions did not exist, the appearance could not exist either. One could avoid this conclusion by denying that appearances are really conditioned, but Kant is deeply committed to appearances being conditioned in this way (since appearances are, e.g., thoroughly causally determined, according to the Second Analogy, and, as we have seen above, causality is a paradigm case of a real conditioning relation).

²⁰ Though this paragraph puts the argument in terms of real conditions, it can be formulated for either logical or real conditions.

²¹ The precise nature of this rational ‘commitment’ is controversial. For discussion of this and related issues, see Chignell (2007). I am sympathetic to calling the appropriate propositional attitude ‘belief’ in Kant’s technical sense, because that term comes as close as any that he recognizes.

²² See also Kant’s remarks about ‘a Something in general’ at A355.

²³ Kant was clearly aware that the argument would have to make further assumptions. See, e.g., 28:272-3.

²⁴ Ameriks’ discussion of this issue in (2000 esp. pp. 47-64) is extremely helpful. Unfortunately, space considerations prevent a fuller defense of the account of transcendental illusion presupposed above.

²⁵ Again, see Ameriks (2000, xx).

²⁶ Note that this claim is distinct from asserting that the *self* is not given in intuition. Such an assertion is controversial and it is not possible in the current context to sort out the senses in which it is and is not correct.

²⁷ Rosefeldt (2000) develops an interesting and detailed account of the logical subject and its role in the Paralogisms. The interpretation presented above is consistent with much of Rosefeldt's interpretation. However, Rosefeldt does not distinguish between logical and real conditions and also does not see the need to think that the logical subject involves any kind of ontological commitment. For Rosefeldt thinks that the concept of the logical subject has a content (*Inhalt*), but no real extension (*Umfang*) (2000, esp. 76-78 and 81-82), whereas the interpretation presented above requires a commitment to a real extension.

²⁸ Just to be clear, there can (indeed must, for Kant) be something real that underlies the I, but whatever that is, it is not a real condition of the I *in the same respect* as the I is the real condition of its representations. This point is analogous to the idea that our free actions are unconditioned, even if we are created by God. What makes both remarks possible is that something can be conditioned in different respects (i.e., different real conditions condition something conditioned in different ways).

²⁹ Dyck (2013, 132-4) argues that Kant is attacking the assumption that only a simple substance can be the causal ground of thought, emphasizing the notion of causality at issue here, due to Wolff's and his followers' reliance on a conception of the soul as the force that is responsible for representations. The argument presented above distinguishes between causal conditions and other kinds of relations as distinct kinds of real conditioning relations, suggests that the latter are at issue, and holds that the main problem with conditioning

relations in the Paralogisms is not that they do not exist, but rather that we cannot cognize them. Wunderlich (2001, 175-88) argues, convincingly, that it is quite unclear how the categories are supposed to fit in both with Kant's statements about the Paralogisms and with the actual subject matter of the Paralogisms (177-79). Thus, though the Second Paralogism is supposed to be about the category of reality and the interpretation here focuses on constitution (as does Kant at, e.g., B420), issues that are distinct, Kant is not always entirely clear on the correspondence and that reality and constitution are closely enough related for this to be a plausible interpretation.

³⁰ Now it is true that what instantiates the logical subject might be a conditioned entity, but it would be conditioned in some respect other than the respect in which the logical subject is unconditioned. See note 28 above.

³¹ It is also compatible with Kant's famous remark that "through this I, or He, or It (the thing), which thinks, nothing further is represented than a transcendental subject of thoughts = x, which is recognized only through the thoughts that are its predicates" (A346/B404). For one must distinguish between the representational content of the logical subject, those features of the object to which the logical subject refers, and the representation by means of which one can represent and refer to the logical subject.

³² It is striking that Kant contrasts 'the logical unity of the subject (simplicity)', with 'the real simplicity of my subject', noting that the concept of a substance 'is valid only of the condition of our cognition, but not of any particular object that is to be specified' (A356).

³³ Note that this argument holds independently of Kant's stance on what Ameriks refers to as appearance immaterialism and scientific immaterialism (cf. Ameriks 2000, 32-42). Phenomenal immaterialism is a separate and quite

complicated issue (cf. Ameriks 2000, 42-5). At A379-80, Kant seems to accept dualism at the phenomenal level, though his argument in favor of it could appear problematic. Allison (1996, 92-106) discusses the relevance of the unity of apperception to scientific materialism (in the context of a discussion of Paul Churchland's views).

³⁴ Again, see Ameriks (2000, 37-42) for the most sophisticated discussion of this issue.

³⁵ I thank audience members at the conference on Early Modern Materialism, organized by Falk Wunderlich at Mainz University in 2014, for helpful comments on a much earlier version of this paper. I also thank Lucy Allais, Karl Ameriks, Rosalind Chaplin, Corey Dyck, Max Edwards, Stefanie Grüne, Tobias Rosefeldt, Joseph Stratmann, Clinton Tolley, Marcus Willaschek, Falk Wunderlich, and an anonymous referee for this journal for extremely helpful comments on a later version of this paper. *If* I am a person, then I am responsible for any remaining errors.